Advanced Mathematical Economics

Paulo B. Brito
PhD in Economics: 2022-2023
ISEG
Universidade de Lisboa
pbrito@iseg.ulisboa.pt

Introduction
15.9.2022

Contents

Chapter 1

Introduction

This is a course on **functional equations** applied to Economics. A functional equation is an equation, with a known form, defined over unknown functions. Therefore the solution of a functional equation, if it exits, is a function.

We introduce two restrictions: first, the domain of the function (or functions) we seek to determine is the set of real numbers (and vectors), and, second, the form of the equation involves a variational mechanism. Usually we are interested in finding or characterising the behaviour of that function which represents the a distribution, or the state of some evolving system.

The set of real numbers has, at least, two features that interest us in this work: it is "continuous" set [1] and its an ordered set.

The main functions we are interested represent aggregate or individual economic variables that represent quantities (as consumption, production, investment, income, asset positions, population, labour effort, employment) and price or returns (as interest and wage rates).

The main independent variable, at least in macroeconomics and growth theory, is time. But we can also (or instead) have space, characteristics of agents (age, skills), dimension (measured by the level of one or more stock variables), or the states of nature (in the probabilistic sense). In spatial dynamics, heterogeneous agents, distributional models we can have more than one independent variable, time and space, skills, income.

We require that those independent variables belong to a continuous domain defined to have an order relationship. The existence of an order relationship means that the independent variable is more than an indexing device for modelling heterogeneity. Its value involves variations along present-future, nearby-faraway, high-low skill, poor-rich, small-large economic units, for example.

The dependent variable also belongs to a continuous set of functions, that we call state space.

There are other types of indexing without an order structure. In this case we can say we have a set of varieties rather than a set of variations, as in our case. We can deal with a variety case by increasing the dimension of the dimension of the state space, i.e, the number of the dependent functions in our model.

The mapping (or mappings) we are trying to find, or at least characterise, is an application

[1] Although this is mathematically ambiguous, see continuous set interpretation.

between two continuous sets. We adopt a variational perspective on the mechanism that generates dynamics or distributions: local interaction of the dependent variable is specified by its gradient. This means that some type of calculus is involved, [2] In some case, the dynamics for a single point in the system (or for a single agent) depends on the evolution of the dynamics throught the whole state space.

Furthermore, having the fundamental dynamic mechanism defined over local interactions does not mean that the dynamics is differentiable or even homogeneous throughout all the state space. If there is some form of non-linearity we can distinguish between local dynamics and global dynamics, small variations and large variations.

Adopting a variational approach allows for importing from (and hopefully exporting as well) an immense pool of insights, theories, methods, and results from mathematics and a wide span of applied sciences, and also allows for seeing more clearly what are the particularities one encounters in setting up and solving economic models. Although the particularities of the main variables may be different, for different topics, the models share some common formal structures. In particular, we will consider models defined by ordinary differential equations (ODE) , partial differential equations (PDE), first-order and parabolic, and stochastic differential equations (SDE), and the related optimization (optimal control of ODEs, PDEs and SDEs).

1.1 Economics in the continuum and economic dynamics in continuous time

The applications in Economics cover a wide range of subjects: dynamic macroeconomics, growth theory, dynamic microeconomics, environmental economics, population economics, international economics, spatial economics, microeconomics (specially mechanism design and matching), finance and possibly all fields of economic theory.

Most economic models in the continuum are related to macroeconomics and growth theory and feature time as the independent variable. In mathematical finance the major analytical contributions are modelled jointly in continuous-time and in a continuous probability space, allowing to study random dynamics. Macroeconomic dynamic general equilibrium models with heterogeneous agents and dynamic game theory take both time and another continuous variable as independent variables allowing to study the dynamics of distributions, i.e, the time-varying behavior of distributions of capital and income. Important theoretical developments in geographical economics or contract theory for example, consider a continuum of space or of types of agents within a static, i.e., time-independent domain. However, increasingly these fields are joining macroeconomics with heterogeneous agents in featuring the dynamics jointly in a continuous domain for space, types, income, capital and continuous time.

Continuous-time dynamic models have been at the core of growth theory and mathematical

[2] For the history and the fundamental role of calculus in the history of science and mathematics see Strogatz (2019).

finance and until the late 1980s in macroeconomics[3]. Since the 1990s dynamic general equilibrium (DGE) and dynamic stochastic general equilibrium models (DSGE) became the dominant paradigm in macroeconomics. Most DGE and DSGE models have been framed in discrete time (see Ljungqvist and Sargent (2012) and Miao (2014)). This allows for a quantitative calibration, simulation and estimation. However, the cost is introducing too much detail associated with the timing of the decisions and renders the qualitative analysis of the models more difficult.

Recently, the research in macroeconomics turned to DGE and DSGE models with heterogeneous agents. This is starting to imply a comeback of continuous-time modelling (see Brunnermeier and Sannikov (2016) and Gabaix et al. (2016) and the references therein.)

In other areas in economic theory modelling in the continuum was still dominant: in spatial economics (see Fujita and Thisse (2002)), in finance (see Cvitanić and Zapatero (2004) and Stokey (2009)). Microeconomics, in which most research features static models, is also increasing to address dynamics, in contracts (see Cvitanić and Zhang (2013)), search in labor markets and networks, just to name a few.

There are several advantages of using a continuous-time framework[4]: first, obtaining qualitative dynamics results, in particular asymptotic dynamics, is not only easier but can also be done drawing on a large body of results from other disciplines (applied mathematics, physics and mathematical biology); second, extending aggregate and/or deterministic models by including dynamics of distributions and stochastic dynamics can be done analytically; and the possibility of obtaining closed form solutions or deriving qualitative dynamic results is in more likely than for the analogous discrete-time and space models.

One important clarification to make is related to the nature of the long-run, or the asymptotic state in the models. The conceptual and the mathematical significance of what sometimes is called "the long-run", be it time dependent or time-independent differ. While mathematically it usually means that the system has reached a state in which there is permanent replications (a fixed point or a balanced growth path), or an asymptotic state, conceptually it is related to the time frame in which the real processes we are trying to model take place.

For instance, the separation between dynamic macroeconomics and growth economics is related to adjustment in the economy that take place at the business-cycle frequency while growth economics addresses long-run adjustments of the major capital stock variables, like the physical capital stock, human capital stock, but also environmental capital. However, both types of dynamic models can display transitional dynamics towards the particular steady state (or balanced growth) variables which concern them. Therefore, while mathematically the two models can display convergence to a mathematical long-run, conceptually those different "long-runs" correspond to real processes with different types of permanence. From this point of view, the long-run in most economic growth models is conceptually different from the long-run in environmental growth models.

---

[3] The state of the art at that time can be seen in v.g. Burmeister and Dobell (1970), Turnovsky (1977), or in Gandolfo (1997)

[4] Discussions continuous versus discrete time modelling Isohätälä et al. (2016) Brunnermeier and Sannikov (2016).

1.2 Generic structure of the models

Next we present a general structure of the equations we will dealing with for our variational approach in the continuum: a general partial differential equation.

1.2.1 Sets, mappings and equations

But before, in a heuristic way, we present the most basic structures we will be dealing next: sets, mappings and functional equations.

A **set**, or **space**, is a collection of objects, which, for our purposes, can be numbers, vectors, Borel sets, functions or operators. Without dealing with the particular topologies, there are two important characteristics which will concern us: convexity and openness/closedness.

Among the sets we will consider are the set of real numbers, of dimension one or higher, i.e. $\mathbb{R}$ or $\mathbb{R}^n$ for $n > 1$, which are sets endowed with a Euclidean metric (normed linear spaces). We will also consider measurable spaces which are spaces of Borel sets over a space of events.

A **function**, f, is a mapping between two numbers, or a mapping between two number spaces, such that $y = f(x)$ where $x \in X \subseteq \mathbb{R}^m$ and $y \in Y \subseteq \mathbb{R}^n$. A function space $\mathcal{F}$ can also be defined as a collection of functions. There are two important characteristics that will concern us: continuity and differentiability.

A **functional**, F is a mapping between a function and a number such that $y = \mathsf{F}[f]$ where $f \in \mathcal{F}$ and $y \in Y$. Most functionals in economics are linear, such as $\mathsf{F}[f] = \int F(f(x))dx$, but we can thing of multiplicative functionals such as $\mathsf{F}[f] = \exp\{\int \ln F(f(x))dx\}$.

An **operator**, $\mathcal{F}$, is a mapping between two functions such that $g = \mathcal{F}\{f\}$, where $g \in \mathcal{G}$ and $f \in \mathcal{F}$.

An **algebraic equation** can be defined by $f(x) = 0$, where $x \in X$. Solving an algebraic equation means finding, at least, one element of X, denoted by x^*, that is a number such that $f(x^*) = 0$. Therefore, algebraic equations are equations whose solutions are numbers.

A **functional equation** is an equation whose solutions are functions. One particular type of functional equations are ordinary and partial differential equations we will be dealing in this course.

1.2.2 Differential equations

The functional equations we deal in this course are differential equations

Consider an (unknown) function $\mathbf{y}(\mathbf{z})$, defined over an independent $\mathbf{z}$, mapping

$$\mathbf{y} : Z \to Y,\ Y \in \mathbb{R}^n$$

where Z may have different types of topology (for instance, it can be an Euclidean space or a probability space) but satisfies $\dim(Z) = m$. Therefore, $\mathbf{z} = (z_1, \dots, z_m)$ and

$$\mathbf{y} = \mathbf{y}(\mathbf{z}) = \begin{pmatrix} y_1(\mathbf{z}) \\ \dots \\ y_n(\mathbf{z}) \end{pmatrix} = \begin{pmatrix} y_1(z_1, \dots, z_m) \\ \dots \\ y_n(z_1, \dots, z_m) \end{pmatrix}.$$

A **differential equation** is a functional equation of the form

$$\mathbf{F}(D^p_{\mathbf{z}}(\mathbf{y}), \dots, D_{\mathbf{z}}(\mathbf{y}), \mathbf{y}(\mathbf{z}), \mathbf{z}) = 0 \tag{1.1}$$

where the $\mathbf{F}(.)$ is known, where $D_{\mathbf{z}}(\mathbf{y})$ is an appropriately defined gradient

$$D_{\mathbf{z}}(\mathbf{y}) \equiv \left(D_{z_1}(\mathbf{y}), \dots, D_{z_m}(\mathbf{y})\right)$$

where

$$D_{z_i}(\mathbf{y}) \equiv \begin{pmatrix} \frac{\partial y_1}{\partial z_i} \\ \dots \\ \frac{\partial y_n}{\partial z_i} \end{pmatrix}, \text{ for } i \in \{1, \dots, m\}$$

and $D^p_{\mathbf{z}}(\mathbf{y})$ is the multi-dimensional matrix of higher-order derivatives.

In most economic applications the equations also involve a vector of parameters $\boldsymbol{\varphi} \in \Phi \subseteq \mathbb{R}^q$ then equation (1.1) becomes

$$\mathbf{F}(D^p_{\mathbf{z}}(\mathbf{y}), \dots, D_{\mathbf{z}}(\mathbf{y}), \mathbf{y}(\mathbf{z}, \boldsymbol{\varphi}), \mathbf{z}, \boldsymbol{\varphi}) = 0 \tag{1.2}$$

There are four numbers to recall, n, m, p, and q, the dimension of the set of dependent variables, the dimension of the set of independent variables, the maximum order of differentiation of the differential equation (1.1), and the dimension of the parameter space in (1.2).

Solving a differential equation means finding at least one function $\boldsymbol{\phi}(\mathbf{z})$, mapping $\phi : Z \to Y \subseteq \mathbb{R}^n$, that satisfies equation (1.1). That is, when substituted in equation (1.1), it satisfies

$$\mathbf{F}(D^p_{\mathbf{z}}(\boldsymbol{\phi}), \dots, D_{\mathbf{z}}(\boldsymbol{\phi}), \boldsymbol{\phi}(\mathbf{z}), \mathbf{z}) = 0.$$

We call it **general solution** of equation (1.1). Function $\phi(\mathbf{z})$, if it exists, may be unique or multiple. Finding a solution means knowing, or at least characterising $\mathbf{y}(\mathbf{z})$ as $\mathbf{y} = \boldsymbol{\phi}(\mathbf{z})$

Different types of differential equations are obtained for different sets of independent variables, Z and m, for different n, p and for different properties of function $\mathbf{F}(.)$.

We see some of examples that we will be dealt in the course in the next sections.

1.2.3 Problems

We will see that, if it exists, a solution to a differential equation represent a generic process or distribution, or evolving distribution.

In most applications, we have problems, which consist in a differential equation together with **additional constraints** on the state of the system. This additional information is related to a direct or observed knowledge of the state of the system for particular value of the independent variables, or with some constraints on the state of the system.

A problem is well posed if the general solution of the differential equation is consistent with the verification of those additional constrains.

1.2.4 Models

Models involving differential equations can be specified directly from a variational principle, or from maximisation of a functional. In the lat case we usually have optimal control problems.

In the next section we specify the main types of models we will be dealing with.

1.3 Differential equations

The main types of equations we will be dealing with are: ordinary differential equations, partial differential equations, and stochastic differential equations.

1.3.1 Ordinary differential equations

Ordinary differential equation (ODE) are differential equations in which the independent variable is of dimension one and belongs to a subset of the set of real numbers: that is $m = 1$, $\mathbf{z} = z \in Z$ and $Z \equiv (z_0, z_1) \subseteq \mathbb{R}$. Therefore, (1.1) becomes

$$\mathbf{F}(D^p(\mathbf{y}), \dots, D\ (\mathbf{y}), \mathbf{y}(z), z) = 0.$$

The value of p gives the order of the equation. If $p = 1$ the equation is called **first order equation**, if $p = 2$ it is called **second order equation**, and so forth. However, all equations with $p \geq 2$ can be transformed into first order equations by defining the derivatives of $y(.)$ as new variables.

This means the an ODE can have the general representation

$$\mathbf{F}(D(\mathbf{y}), \mathbf{y}(z), z) = 0 \tag{1.3}$$

where, again, $\mathbf{F}$ is known, and

$$D(\mathbf{y}) = \left(\frac{dy_1(z)}{dz}, \dots, \frac{dy_n(z)}{dz}\right)$$

is the gradient as regards the independent variable z. Equation (1.3) is called **implicit ODE**.

Let the gradient of $\mathbf{F}$ as regards $D(\mathbf{y})$ be

$$\mathbf{A}\ (\mathbf{y}, z) = D_{D(\mathbf{y})}\mathbf{F}.$$

If $\mathbf{F}_{D(\mathbf{y})}(.)$ is monotonic and regular we can transform, at least locally, equation (1.3) into

$$\mathbf{A}\ (\mathbf{y}, z)D(\mathbf{y}) + \mathbf{G}(\mathbf{y}, z) = 0 \tag{1.4}$$

where $\mathbf{A}(.)$ is a $n \times n$ matrix and $\mathbf{G}(.)$ is a $n \times 1$ vector.

If function $\mathbf{F}$ is separable in $D(\mathbf{y})$, as in (1.4), we call it **quasi-linear equation**, because it is linear in the derivatives. This allows us to write equation (1.4) in the **explicit form**

$$D(\mathbf{y}) = \mathbf{H}(\mathbf{y}, z). \tag{1.5}$$

Intuitively, if $\mathbf{H}$ is linear, we can see equation (1.5) as describing a movement over a Euclidean space, while equation (1.3) describes a movement over a generalized surface (which can be regular or not).

The solution of an ordinary differential equation, in its implicit form (1.3) or in its normal form (1.5), is a function $\phi : Z \to Y \subseteq \mathbb{R}^n$, called **general solution**, such that $\mathbf{y} = \boldsymbol{\phi}(z)$ solves the differential equation.

Now, we can be more specific about the additional information we referred to in the previous section. Let, again, $z \in [z_0, z_1]$, where $z_0 < z_1$. If we set $\mathbf{y}(z_0) = \mathbf{y}_0$, which is a known element of set Y, we call

$$\begin{cases} \mathbf{F}(D(\mathbf{y}), \mathbf{y}(z), z) = \mathbf{0}, \\ \mathbf{y}(z_0) = \mathbf{y}_0 \end{cases}$$

an **initial-value problem**. The solution of this problem, $\mathbf{y} = \boldsymbol{\phi}\ (z, \mathbf{y}_0)$, is called **particular solution** and it is a function of the independent variable and the data on the system $\mathbf{y}_0$. Observe that the dimension of $\mathbf{y}_0$ is the same as the number of equations in $\mathbf{F}(.) = \mathbf{0}$. If we keep the same relationship between the number equations, (in the ODE system) but some variables are fixed at the z_1 we say we have a **boundary-value** problem.

Next we present some low-dimensional ODE models and related concepts we will deal further in the course.

Scalar ODE

In this subsection we set $n = 1$ and consider again the one-dimensional independent variable $z \in Z \subseteq \mathbb{R}_+$ and let $Y = \mathbb{R}$. When convenient we set $Z = [z_0, z_1]$ (or $[z_0, z_1)$, $(z_0, z_1]$ or (z_0, z_1)).

To obtain an intuition on the variational approach to dynamics we consider the setting up of a particular scalar ODE. Let $y(s + h) \in Y$ be the value of y for $z = s + h$ and let $y(s) \in Y$ be the value of y for $z = s$. Assume that the **variation** of y depends on h, that is on the variation of z, such that

$$y(s + h) - y(s) = f(y(s), s)(s + h - s) = f(y(s), s)h,$$

then

$$\frac{y(s + h) - y(s)}{h} = f(y(s), s)$$

Letting h be very small, and because the derivative of $y(z)$, taken at the point $z = s$, is

$$\frac{dy(s)}{ds} = \lim_{h \to 0} \frac{y(s + h) - y(s)}{h},$$

if a derivative exists for all $z \in Z$, we obtain a **scalar ODE** in the normal form,

$$y^{'}(z) \equiv \frac{dy(z)}{dz} = f(y(z), z).$$

Therefore, the ODE represents the change of a variable, y, if there is an infinitesimal variation of the independent variable, given the initial level for both the dependent and the independent variables (this is the reason why we need an order structure imposed on Z).

If the **independent variable is time**, and $n = 1$ and $z = t \in \mathrm{T} \subseteq \mathbb{R}_+$ where T is the time interval, usually $\mathrm{T} = [0, T]$ with T finite or $\mathrm{T} = [0, \infty)$. In this case the function $y : \mathrm{T} \to \mathrm{Y} \subseteq \mathbb{R}$ and $D(y) = y'(t)$ is denoted by $\dot{y}$, and the derivative is [5]

$$\dot{y} \equiv \frac{dy(t)}{dt} = \lim_{\epsilon \to 0} \frac{y(t+\epsilon) - y(t)}{\epsilon}.$$

The implicit ODE takes the form

$$F(\dot{y}, y, t) = 0.$$

Solving an ODE means finding (or proving the existence, multiplicity and characterising) a function $y(t) = \phi(t)$ such that

$$F(\dot{\phi}(t), \phi(t), t) = 0.$$

Properties of $F(.)$ The solution of the differential equation (roughly) inherits the properties of $f(.)$, as regards continuity and differentiability. We can consider the following cases:

- if $F_{\dot{y}}(\dot{y}, y)$ is non-differentiable as regards $\dot{y}$. Consider the quasi-linear equation

  $$a(y)\dot{y} = g(y)$$

  where there are point y_s such that $a(y_s) = 0$ then $f(.)$ is **not regular** in equation

  $$\dot{y} = f(y) \equiv a(y)^{-1} g(y)$$

  which means that it is not locally Lipschitz (i.e. $\lim_{y \to y_s} \dot{y} = \pm\infty$). In this case we have a **constrained** ODEs or **singular** ODEs.

- if $F_{\dot{y}}(\dot{y}, y)$ is differentiable as regards $\dot{y}$ but $f(.)$ is locally non-differentiable or non-continuous we can write the ODE in normal form, as

  $$\dot{y} = \begin{cases} f_1(y) & \text{if } h(y) \leq 0 \\ f_2(y) & \text{if } h(y) > 0 \end{cases}$$

  where, if $y = \tilde{y}$ such that $h(\tilde{y}) = 0$. We can have two cases

  - **non-differentiable**: $f_1(\tilde{y}) = f_2(\tilde{y})$ and $D(f_1(\tilde{y})) \neq D(f_2(\tilde{y}))$
  - **non-continuous**: $f_1(\tilde{y}) \neq f_2(\tilde{y})$ and $D(f_1(\tilde{y})) \neq D(f_2(\tilde{y}))$

- both $F_{\dot{y}}(\dot{y}, y)$ is differentiable as regards $\dot{y}$ and $f(.)$ are **continuous, differentiable and regular**. In this case the ODE takes the form

  $$\dot{y} = f(y)$$

[5] We will use the dot notation, $\dot{y}$, for time derivatives and the $y^{'}(z)$ notation for non-time independent variables.

The quasi-linear equation can be written as

$$a(y,t)\dot{y} + g(y,t) = 0$$

If $a(.)$ is everywhere different from zero, we say have the **non-autonomous** ODE in the normal form [6] and write

$$\dot{y} = \frac{dy}{dt} = f(y,t).$$

If time does not enter explicitly as an argument of $f(.)$, we say we have an **autonomous** ODE (in the normal form)

$$\dot{y} = \frac{dy}{dt} = f(y). \tag{1.6}$$

In most applications we consider a set of parameters $\varphi \in \Phi$

$$\dot{y} = f(y, \varphi).$$

ODEs and problems involving ODEs Let $y = y(t)$ for $t \in [0, T]$

Initial value problem: defined by an ODE and an initial value for the unknown function

$$\dot{y} = f(y) \text{ and } y(0) = y_0 \text{ known}$$

Boundary value problem: defined by an ODE and a terminal value for the unknown function

$$\dot{y} = f(y) \text{ and } y(T) = y_T \text{ known.}$$

Integral representation Another intuition can be drawn from the an integral representation of a ODE, where backward or forward dependencies can be more clearly seen.

Let $0 < t < T$ and let $y(t)$, display a dependency of its **past** values as

$$y(t) = \int_0^t f(y(s))ds$$

that is, is the value at t of a function is an integral (a generalized sum) of a function of its past values. It we take a time derivative, and apply the Leibnitz rule, we have equation (1.6):

$$\frac{dy(t)}{dt} = \frac{d}{dt}\Big(\int_0^t f(y(s))ds\Big) = f(y(t)).$$

In the case in which the value of variable at time t, $y(t)$, is a function of its **future** values as

$$y(t) = -\int_t^T f(y(s))ds,$$

[6] The convention of not writing the time-dependence of the dependent variable y is common in the literature.

we get equation (1.6) time-differentiating.

Using a classification which is common in the stochastic differential equations (SDE) literature, we can say we have a **forward ODE** when we have a ODE jointly with an initial value, and we can say we have a **backward ODE** when we have a ODE jointly with a terminal value.

Integral equations and ordinary differential equations While an ODE is a functional equation, an integral equation is an operator equations. For instance

$$y(t) = \int_X K(t,s) f(y(s)) ds$$

is an integral equation. Like the ODE an integral equation solution is a function, however, while the differential equation involves a local dependency the integral equation involves a dependency of a global nature.

Examples of scalar ODEs

Example 1: the exponential growth model for population growth

$$\dot{N} = \mu N$$

where $N(t)$ is population at time t and μ is a parameter representing the instantaneous rate of growth of the population. This ODE is usually interpreted as a forward equation, i.e, we have information on $N(0)$ and want to find the behavior os population $N(t)$ for $t > 0$.

Example 2: a generic budget constraint

$$\dot{W} = Y(t) - D(t) + r(t)W$$

where W is the stock of financial assets, Y and D are non-financial income and expenditures, and r, is the instantaneous rate of return. This equation can be interpreted as a forward or a backward equation. In the first case we know $W(0)$ and want to determine the future behavior of $W(t)$ and in the second case we fix a, usually bounded, value for $W(T)$ (or for its present value) and want to determine the initial value $W(0)$ which is consistent with it. The so-called sustainability analysis can be conducted in this way.

Example 3: the Solow growth model

$$\dot{k} = s(k) - \delta k$$

where k is the per capital capital stock, $s(k)$ is the savings function, and δ is the rate of depreciation of capital. This ODE is also usually interpreted as a forward ODE.

Planar and higher-dimensional ODE

If $n = 2$, and keeping $z = t$, we have the **planar** ODE, where $\mathbf{y}(t) = (y_1(t), y_2(t)) \in Y \subseteq \mathbb{R}^2$. The ODE in implicit

$$F_1(\dot{y}_1, \dot{y}_2, y_1, y_2) = 0$$
$$F_2(\dot{y}_1, \dot{y}_2, y_1, y_2) = 0$$

where $F_i(.)$, for $i = 1, 2$ can take non continuous, or non-differentiable regular or singular forms as for the planar equation. If $F(.)$ is well-behaved we have the ODE in its normal form

$$\begin{aligned} \dot{y}_1 &= f_1(y_1, y_2) \\ \dot{y}_2 &= f_2(y_1, y_2). \end{aligned}$$

In this case we can also have the initial value (or the forward ODE) and the terminal-value problems (or the backward ODE) cases as in the scalar case. However, we have a new case:

Mixed boundary-initial value problem: defined by an ODE and a number of initial and terminal value conditions which is equal to the dimension of y (n). Example: let $y = (y_1, y_2)$

$$\begin{aligned} \dot{y}_1 &= f_1(y_1, y_2),\ y_1(0) = y_0 \\ \dot{y}_2 &= f_2(y_1, y_2),\ y_2(T) = y_T \end{aligned}$$

The optimality condition for optimal control problems take the form (for $T \to \infty$)

$$\begin{aligned} \dot{y}_1 &= f_1(y_1, y_2),\ y_1(0) = y_0 \\ \dot{y}_2 &= f_2(y_1, y_2),\ \lim_{t\to\infty} h(y_1(t), y_2(t), t) = 0 \end{aligned}$$

We can say that in this case we have a **forward-backward ODE**

Examples of planar ODEs **Example 4**: the Ramsey model featuring the coupled dynamics of consumption, c and capital, k, and is

$$\begin{aligned} \dot{c} &= c(r(k) - \rho) \\ \dot{k} &= y(k) - c \end{aligned}$$

the first equation has been called several names, such as Euler equation, Keynes-Ramsey rule, consumer arbitrage condition, and the second is a budget constraint.

Example 5: the Ramsey model with endogenous labor

$$\begin{aligned} \dot{c} &= c(r(k, l) - \rho) \\ \dot{k} &= y(k, l) - c \\ c &= C(k, l) \end{aligned}$$

where l is the labor effort can be transformed into a planar equation.

Those are both cases of forward-backward ODEs.

Higher-dimensional ODE's

If $n > 2$ the ODE in normal form is

$$\begin{aligned} \dot{y}_1 &= f_1(y_1, \dots, y_n) \\ &\dots \\ \dot{y}_n &= f_n(y_1, \dots, y_n) \end{aligned}$$

we have a **multidimensional** ODE.

Example 6: the Ramsey model with government debt dynamics can have the form

$$\begin{aligned}\dot{c} &= c((1-\tau)r(k,b)-\rho)\\ \dot{k} &= y(k)-c\\ \dot{b} &= g-\tau y(k)+r(k,b)b\end{aligned}$$

where b is the level of government debt, τ is the tax rate, and g public expenditures. Again this is a forward-backward ODEs, where the dimension of the forward (backward) component depends on the condition on b.

Solving ODEs and problems involving ODEs

Ideally we would like to find function $\phi(t,.)$ explicitly. However this is only possible for a small number of equations.

The three main issues regarding solving differential equations that cannot be solved explicitly are associated with:

- existence of solutions
- uniqueness of solutions
- characterization of solutions, i.e., their behavior relative to the independent variable, t, and the other data of the problem (parameters, initial or terminal values). The main tool for this is the **qualitative theory of ODE** or **bifurcation theory**.

1.3.2 Partial differential equations

A **partial differential equation** (PDE) is one equation involving one or more than one function of at least two independent variables together with its derivatives. In economics applications one of the independent variable is time. In equation (1.1) we have $m > 1$, $Z \subseteq \mathbb{R}^m$, $p = 2$ and $\mathbf{F}(.)$ be monotonic and regular in all the derivative arguments.

First-order PDE

Let $m = 2$ and consider the function $\mathbf{y} = \mathbf{y}(t,x)$ where $(t,x) \in \mathbb{R}^2$.

A **first-order partial differential equation** takes the general form

$$\mathbf{F}(D_t(\mathbf{y}), D_x(\mathbf{y}), \mathbf{y}(t,x), t, x) = 0$$

If $n = 1$ a **quasi-linear** hyperbolic PDE is

$$a(t,x)y_t + b(t,x)y_x = f(y),$$

where $y : \mathbb{R}^2 \to Y \subseteq \mathbb{R}$, $y_t = \frac{\partial y(t,x)}{\partial t}$ and $y_x = \frac{\partial y(t,x)}{\partial x}$. Some times we represent the partial derivatives by $\partial_t y(t,x)$ and $\partial_x y(t,x)$.

This type of equations models for instance transport, conservative, age-dependent distributions along time. It represents a distribution moving along time.

Example 6 Dynamics of an age-dependent population. The density of population $n = n(a, t)$ is equal to the number of individuals of age a at time t is governed by the McKendrick PDE

$$n_t + n_a = \mu(a)n$$

where $\mu(.)$ is the mortality rate. In general the equation is complemented with a condition for the newborns $n(0, t) = \int_0^{a_{\max}} \beta(a)n(a, t)da$ when fertility is age-dependent.

Parabolic partial differential equations

A parabolic PDE has the general form

$$\mathbf{F}(D_x^2(\mathbf{y}), D_t(\mathbf{y}), D_x(\mathbf{y}), \mathbf{y}(t, x), t, x) = 0$$

that is, it involves a second derivative as regards the "spatial" variable.

If $n = 1$ a **quasi-linear** parabolic PDE is

$$a(t, x)y_t + b(t, x)y_x + c(t, x)y_{xx} = f(y),$$

where $y : \mathbb{R}^2 \to Y \subseteq \mathbb{R}$ and $y_{xx} = \frac{\partial^2 y}{\partial x^2}$.

In Economics and Finance there are two types of PDE's we should distinguish:

- **forward** parabolic PDE
$$y_t(x, t) - y_{xx}(x, t)) = f(y(x, t))$$
models forward diffusion phenomena starting from one known initial distribution $y(0, x) = \phi(x)$ and diffusing out through time;

- **backward** parabolic PDE
$$y_t(x, t) + y_{xx}(x, t)) = f(y(x, t))$$
is very common on mathematical finance, in which a terminal distribution $y(T, x) = h(x)$ is known and an initial distribution $y(0, x) = \phi(x)$ is to be determined.

Example 7 The heath equation is the simplest case of a forward equation

$$u_t(x, t) = u_{xx}(x, t)$$

where $u(x, t)$ is the temperature of an one-dimensional rod, at location x at time t.

Example 8 The well known Black-Scholes is an example of a backward equation

$$v_t(S, t) = -\frac{\sigma S^2}{2} v_{SS}(S, t) + (v(S, t) - S v_S(S, t))$$

where $v(S, t)$ is the value of an option over an underlying asset with price S at time t, σ is the instantaneous volatility of the underlying asset and r is the risk-free interest rate.

1.3.3 Stochastic differential equations

Let: $m = 2$, $Z \subseteq \mathbb{R} \times \Omega$ (Ω is a probability space) , $p = 1$ and $F(.)$ monotonic and regular in all the derivative arguments. A stochastic differential equation is defined over $Y = Y(t, \omega)$ where $\omega \in (\Omega, \mathsf{F}, \mathbb{P})$ is a probability space. Therefore, Y for a particular t is a random variable and for particular pair (t, ω_0), $Y(t, \omega_0) = y_0(t)$ it is a realization.

As, for every t, $Y(t, \omega)$ is in general not differentiable in the classic sense, we need a theory of differentiation (or integration) of stochastic processes. The most widespread theory of integration of stochastic processes leads to the Itô stochastic differential equation.

The Itô's stochastic differential equation A **Itô's stochastic differential equation** (SDE) takes the form

$$dy = f(y)dt + \sigma(y)dW$$

where dW is a standard Wiener process, i.e., a stochastic process following a normal distribution with zero mean and variance dt: $dW \sim N(0, dt)$. An important fact about $y(t, \omega)$ is that it is not differentiable (in the classic sense) as regards t. Therefore, in order to solve a SDE we need to apply the **Itô's or stochastic calculus**.

There are relationships between SDE's and PDE's. Let $p(t, y) = \mathbb{P}[\ Y(t)|Y(0) = 0]$, i.e, the probability distribution that $Y(t, \omega) = y$, when as of time $t = 0$ $Y(0) = 0$. It can be shows that it satisfies the Kolmogorov forward or Fokker-Planck equation

$$\partial_t p\ (t, y) = \frac{1}{2}\ \partial_{yy}[\sigma(y)p(t, y)]\ - \partial_y[f(y(t, y)p(t, y)]$$

which is a parabolic PDE.

1.4 Functionals and optimization with ODE's

We saw that a functional is a mapping between a function space and a number space. Let $f : X \to Y$, where $X, Y \in \mathbb{R}$ for and consider the functional $\mathsf{F}[f]$. The usual definition of derivative cannot be applied here because f is an infinite-dimensional object.

In order to perform variational analysis with functionals we need a new concept for derivatives in infinite-dimensional spaces.

Consider a perturbation of function f along the direction $\delta f(x)$ such that $f(x)$ is perturbed to $f(x) + \epsilon\delta f(x)$. The variation of the functional is $\mathsf{F}[f + \epsilon\delta f] - \mathsf{F}[f]$. A **Gâteaux** derivative is

$$\delta\mathsf{F}[f]_v = \lim_{\epsilon\to 0} \frac{\mathsf{F}[f + \epsilon\delta f] - \mathsf{F}[f]}{\epsilon}.$$

Consider the linear functional

$$\mathsf{V}[y]\ = \int_X h(y(x))dx.$$

Then, the Gâteaux derivative of this functional, along the direction $v(x)$, is

$$\delta \mathsf{V}[y]_v = \int_X h^{'}(y(x))\, v(x) dx$$

In order to prove this ...

Problems

$$\max_y \int_X h(y(x))dx.$$

First order condition

$$\delta \mathsf{V}[y]_v = 0$$

for any $v(x)$. Therefore, we obtain a functional equation

$$h^{'}(y(x)) = 0$$

The optimum, if it exists is a function $y^*(x)$.

If there are constraints

$$\max_y \left\{ \int_X h(y(x))dx : \int_X g(y(x))dx = 0 \right\}$$

the first order conditions are

$$\begin{cases} h^{'}(y(x)) + \lambda g^{'}(y(x)) = 0 \\ \int_X g(y(x))dx = 0 \end{cases}$$

the optimum is a pair: a function $f^*(x)$ and a number λ^* (a Lagrange multiplier).

This type of structure is typical in infinite dimensional static problems in microeconomics.

1.4.1 Optimal control of ODE's

Consider two unknown functions $\mathbf{u} : \mathrm{T} \to \mathbb{R}^u$ and $\mathbf{x} : \mathrm{T} \to \mathbb{R}^n$, and define the **value functional**

$$\mathsf{V}[\mathbf{u}, \mathbf{x}] = \int_{t_0}^{t_1} h(t, \mathbf{u}(t), \mathbf{x}(t))dt \tag{1.7}$$

the ODE

$$\dot{\mathbf{x}} = f(\mathbf{u}, \mathbf{x}). \tag{1.8}$$

and addition conditions on $t = 0$, $t = T$ and the associated values $x(t_0) = x_0$ or $x(t_1) = x_{t_1}$, or restrictions upon them. In the simplest problem, we assume we know $(t_0, \mathbf{x}(t_0)) = (0, \mathbf{x}_0)$ and $(t_1, \mathbf{x}(t_1)) = (T, \mathbf{x}_T)$.

The **optimal control problem** (OCP) is to find the functions $\mathbf{u}^*(.)$ and $\mathbf{x}^*(.)$ that

$$\max_{\mathbf{u}(.)} \mathsf{V}(\mathbf{u}, \mathbf{x})$$

subject to equation (1.8), given $\mathbf{x}(0) = \mathbf{x}_0$ and other information on T or $\mathbf{x}(T)$.

The most common problem in economics has the value function

$$\mathsf{V}(\mathbf{u}, \mathbf{x}) = \int_0^\infty h(\mathbf{u}(t), \mathbf{x}(t)) e^{-\rho t} dt$$

and is called **infinite horizon discounted optimal control problem**.

There are several methods for solving the OCP all leading to a ODE.

From now on let us assume that the control and the state variables are scalar.

Calculus of variations problem

If we can write $\dot{x} = f(u, x)$ as $u = g(\dot{x}, x)$ then the problem becomes

$$\max_{x(.)} \int_0^T F(\dot{x}, x, t) dt$$

The optimality condition is the Euler-Lagrange equation

$$\frac{\partial F(\dot{x}, x, t)}{\partial x} + \frac{d}{dt}\left(\frac{\partial F(\dot{x}, x, t)}{\partial \dot{x}}\right) = 0$$

together with initial and terminal conditions (it is a mixed-value problem).

The EL equation is a second order ODE, evaluated at an optimum,

$$F_x(\dot{x}, x, t) + F_{\dot{x}t}(\dot{x}, x, t) + F_{\dot{x}x}(\dot{x}, x, t)\dot{x} + F_{\dot{x}\dot{x}}(\dot{x}, x, t)\ddot{x} = 0.$$

Pontriyagin's maximum principle

Introducing the co-state variable $q(t)$ and the Hamiltonian function

$$H(x, u, t) = h(x, u, t) + q(t) f(x, u, t)$$

the necessary conditions for an optimum involve the modified Hamiltonian dynamic system (MHDS), which is a system of two first order equations,

$$\begin{aligned} \dot{q} &= -H_x(x, u(q, x), t) \\ \dot{x} &= f(x, u(q, u), t) \end{aligned}$$

if the functions $f(.)$ and $h(.)$ are sufficiently differentiable, together with initial and terminal conditions. Examples 4 and 5 are MHDS

Dynamic programming

Optimality conditions are given by an implicit ODE's the **Hamilton-Jacobi-Bellman** (HJB) equation: value function $V(x,t)$ satisfies at the optimum the HJB equation

$$-V_t(x,t) \;= \max_{u(.)}\{\; h(x,u,t) + V_x(x,t)\; f(x,u,t)\}$$

This is a partial differential equation (PDE) in implicit form.

For infinite-horizon autonomous problems the HJB equation becomes an ODE (in the implicit form)

$$\rho V(x) = \max_{u(.)}\Big\{\; h(x,u) + V^{'}(x)\; f(x,u)\Big\}$$

where $V(x)$ is unknown. The policy function takes the form $u = u(x, V^{'}(x))$ then the HJB function becomes the implicit ODE

$$\rho V(x) = \; h\left(x, u(x, V^{'}(x))\right) + V^{'}(x)\; f\left(x, u(x, V^{'}(x))\right).$$

Extensions

The former definition of optimal control problem can be extended in several different ways, for instance:

- by increasing the number of control variables (see example 5)
- by increasing the dimension of the state vector $x = (x_1, \dots, x_n)$ (which doubles the dimension of the ODE representing the first order conditions)
- by introducing a instantaneous value terminal state or control $V(x(t), u(T), T)$
- by introducing constraints on the terminal state or control $H(x(T), u(T), T) \leq 0$
- by introducing constraints on the trajectories of the state or control variables $H(x(t), u(t), t) \leq 0$.

1.4.2 Optimal control of PDE's

Although much less well known, and usually very hard to solve, there are optimal control problems for systems governed by PDE's. The first order conditions are generally forward-backward PDE's.

Optimal control of first-order PDE's

Consider two unknown functions $u : \mathsf{D} \rightarrow \mathbb{R}^u$ and $y : \mathsf{D} \rightarrow \mathbb{R}^n$ where $\mathsf{D} = (\underline{x}, \bar{x}) \times (0, \infty)$ An **optimal control problem of hyperbolic PDE** has the form

$$\max_{u(.)} \int_{\underline{x}}^{\bar{x}} \int_0^{\infty} h(u(x,t), y(x,t), x, t) dt dx$$

subject to the first-order PDE

$$y_t(x,t) \; + y_x(x,t) = f(u(x,t), y(x,t))$$

plus initial $y(0,x) = y_0(x)$ and possibly boundary conditions. This models the optimal choice of a distribution along time.

The first order conditions, from the **Pontryagin's maximum principle**

Consider the co-state variable $q(t,x)$ and the Hamiltonian function

$$H(h(u(t,x), y(t,x), x, t) = h(u(t,x), y(t,x), x, t) + q(t,x) f(u(t,x), y(t,x))$$

necessary f.o.c. include involve a system of two first-order PDE (one moving forward and the other backward)

$$\begin{aligned} q_t &= -\, q_x + H_y(u(q,y), y, .) \\ y_t &= y_x \; + f(u(q,y), y) \end{aligned}$$

the solution, if it exists, features an optimal distribution evolving along time, possibly converging to a bounded asymptotic distribution.

Parabolic PDE's

Consider two unknown functions $u : Z \rightarrow \mathbb{R}^u$ and $x : Z \rightarrow \mathbb{R}^n$ where $Z = (\underline{x}, \bar{x}) \times (0, \infty)$ An **optimal control problem of parabolic PDE** has the form

$$\max_{u(.)} \int_{\underline{x}}^{\bar{x}} \int_0^{\infty} h(u(x,t), y(x,t), x, t) dt dx$$

subject to the *forward* parabolic PDE

$$y_t(x,t) = y_{xx}(x,t) \; + f(u(x,t), y(x,t))$$

plus initial, $y(0,x) = y_0(x)$, and possibly boundary conditions

The first order conditions, from the **Pontryagin's maximum principle** Consider the co-state variable $q(t,x)$ and the Hamiltonian function

$$H(h(u(t,x), y(t,x), x, t) = h(u(t,x), y(t,x), x, t) + q(t,x) f(u(t,x), y(t,x))$$

necessary f.o.c. include involve a system of two parabolic PDE (one forward and one backward)

$$\begin{aligned} q_t &= -q_{xx} + H_y(u(q,y), y, .) \\ y_t &= y_{xx} + f(u(q,y), y) \end{aligned}$$

the solution, if it exists, features an optimal distribution evolving along time, possibly converging to a bounded asymptotic distribution. Reference: Li and Yong (1995).

1.4.3 Optimal control of SDE's

Consider two unknown functions $u : Z \to \mathbb{R}^u$ and $x : Z \to \mathbb{R}^n$ where $Z = \Omega \times (0, \infty)$ where Ω is again a probability space

An **optimal control problem of SDE** can take the form

$$\max_{u(.)} \mathbb{E}\left[\int_0^\infty h(u(t), y(t), t) dt \right]$$

subject to the SDE

$$dy = f(u, y)dt + \sigma(u, y)dB$$

plus initial and boundary conditions.

Note that

$$\mathbb{E}\left[\int_0^\infty h(u(t), y(t), t) dt \right] = \int_\Omega \int_0^\infty h(u(t,\omega), y(t,\omega), t)\pi(\omega) dt d\omega$$

where $\pi(.)$ is a density function

To find the optimum it is convenient to use the **stochastic dynamic programming principle**

The value function $V(y, t)$ satisfies at the optimum the HJB equation

$$-V_t(y,t) = \max_{u(.)} \left\{ h(y, u, t) + V_y(y, t)\ f(y, u) + \frac{1}{2}\sigma(u, y)^2 V_{yy}(y, t) \right\}$$

which is an implicit parabolic PDE.

Much less known is the SDE version to the maximum principle. Reference: Peng (1990) and Yong and Zhou (1999).

Bibliography

Brunnermeier, M. K. and Sannikov, Y. (2016). Macro, Money and Finance: A Continuous Time Approach. NBER Working Papers 22343, National Bureau of Economic Research, Inc.

Burmeister, E. and Dobell, A. R. (1970). *Mathematical Theories of Economic Growth*. Macmillan.

Cvitanić, J. and Zapatero, F. (2004). *Introduction to the Economics and Mathematics of Financial Markets*. MIT Press.

Cvitanić, J. and Zhang, J. (2013). *Contract Theory in Continuous-Time Models*. Springer Finance. Springer-Verlag Berlin Heidelberg, 1 edition.

Fujita, M. and Thisse, J.-F. (2002). *Economics of Agglomeration*. Cambridge.

Gabaix, X., Lasry, J., Lions, P., and Moll, B. (2016). The Dynamics of Inequality. *Econometrica*, 84:2071–2111.

Gandolfo, G. (1997). *Economic Dynamics*. Springer-Verlag.

Isohätälä, J., Klimenko, N., and Milne, A. (2016). Post-crisis macrofinancial modeling: Continuous time approaches. In Haven, E., Molyneux, P., Wilson, J. O. S., Fedotov, S., and Duygun, M., editors, *The Handbook of Post Crisis Financial Modeling*, pages 235–282. Palgrave Macmillan UK.

Li, X. and Yong, J. (1995). *Optimal Control Theory for Infinite Dimensional Systems*. Systems and Control: Foundations and Applications. Birkhäuser Basel, 1 edition.

Ljungqvist, L. and Sargent, T. J. (2012). *Recursive Macroeconomic Theory*. MIT Press, Cambridge and London, 3rd edition.

Miao, J. (2014). *Economic Dynamics in Discrete Time*. MIT Press.

Peng, S. (1990). A general stochastic maximum principle for optimal control problems. *SIAM Journal on Control and Optimization*, 28.

Stokey, N. L. (2009). *The Economics of Inaction*. Princeton.

Strogatz, S. (2019). *Infinite Powers: The Story of Calculus - The Language of Universe*. Atlantic Books.

Turnovsky, S. (1977). *Macroeconomic Analysis and Stabilization Policy*. Cambridge University Press.

Yong, J. and Zhou, X. Y. (1999). *Stochastic Controls. Hamiltonian Systems and HJB Equations*. Number 43 in Applcations of Mathematics. Stochastic Modelling and Applied Probability. Springer.